



## King's Research Portal

DOI:

[10.1016/j.envsoft.2016.09.006](https://doi.org/10.1016/j.envsoft.2016.09.006)

*Document Version*

Peer reviewed version

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Sun, Z., Lorscheid, I., & Millington, J. D. A. (2016). Simple or complicated agent-based models? A complicated issue. *ENVIRONMENTAL MODELLING AND SOFTWARE*, 86, 56-67.

<https://doi.org/10.1016/j.envsoft.2016.09.006>

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

# Simple or complicated agent-based models? A complicated issue

## Abstract

Agent-based models (ABMs) are increasingly recognized as valuable tools in modelling human-environmental systems, but challenges and critics remain. One pressing challenge in the era of “Big Data” and given the flexibility of representation afforded by ABMs, is identifying the appropriate level of complicatedness in model structure for representing and investigating complex real-world systems. In this paper, we differentiate the concepts of complexity (model behaviour) and complicatedness (model structure), and illustrate the non-linear relationship between them. We then systematically evaluate the trade-offs between simple (often theoretical) models and complicated (often empirically-grounded) models. We propose using pattern-oriented modelling, stepwise approaches, and modular design to guide modellers in reaching an appropriate level of model complicatedness. While ABMs should be constructed as simple as possible but as complicated as necessary to address the predefined research questions, we also warn modellers of the pitfalls and risks of building “mid-level” models mixing stylized and empirical components.

**Keywords:** empirically grounded models; pattern-oriented modelling; stepwise approach; complexity; model complicatedness

# 1 Introduction

Agent-based models (ABMs) have become a well-established approach for studying complex human-environmental systems, such as land-use systems, by explicitly modelling decision-making and dynamic interactions of individuated actors (An, 2012; Filatova et al., 2013; Matthews et al., 2007; O’Sullivan et al., 2015; Parker et al., 2003). ABMs allow modellers to explicitly incorporate feedbacks between human and environmental systems and to investigate emergent patterns at the macro level in time and space due to interactions at lower levels of organization (Batty, 2007). As a result, ABMs continue to gain popularity among modellers.

Part of the success of ABMs stems from their ability to produce emergent system dynamics from often surprisingly simple rules specified at the individual level. The most popular and highly-cited agent based models are rather simple, mainly because they aim at delivering important insights on possible explanations for general patterns (Parker et al., 2003). A famous example is the segregation model by Schelling (1969), based on a simple rule specifying where and under what condition individuals relocate. The model shows how highly segregated patterns of societal groups can result from surprisingly weak aversion of individuals (i.e., relatively high tolerance to another group). Most early ABMs were stylized models to represent general dynamics in, for example, economic systems (Albin and Foley, 1992; Marks, 1992), social systems (Epstein and Axtell, 1996; Schelling, 1969, 1971) and ecological systems (Grimm, 1999; Reynolds, 1987). Often simple theoretical assumptions were made to model agents, partially due to incomplete knowledge of individuals’ interactions or underlying decision mechanisms, but also because of limited data availability at the individual level. According to Parker et al. (2003), such simple models fall into the category of so-called “Picasso”-models—stylized models with a high level of abstraction used to test general principles and ideally yielding generalizable results.

The simplicity of rules and a lack of empirical support mean simple ABMs are often labelled “toy” models, perceived to be suitable only for “proof-of-concept” purposes (Crooks et al., 2008; Janssen and Ostrom, 2006). In contrast, so-called “photograph” models (Parker et al. 2003) depend on empirical data to provide high levels of detail (Balbi et al., 2013). Such models have gained great popularity with increasing numbers of ABMs of land use/land cover change designed and implemented for particular case studies (e.g., O’Sullivan et al. (2012); Piorr et al. (2009)). These empirical ABMs tend to be more complicated and usually demand large amounts of data, defining detailed functions rather than using heuristic rules. The success of empirical ABMs proves that ABMs are not only for “proof-of-concept”, but can also be useful in addressing real-world problems. The emergence of empirical ABMs is partially driven by an increasing demand of stakeholders and decision-makers to provide support for understanding the potential implications of decisions in

complex situations (Smajgl et al., 2011). Moreover, increasing availability of individual-level data (e.g. Poppy et al., 2014; Richards et al., 2014) and greater computing power has accommodated and fuelled the trend.

As a result, we observe an ongoing trend of ABMs becoming more complicated (Janssen and Ostrom, 2006; Rounsevell et al., 2012). The development towards more complicated ABMs however comes with several challenges: complicated ABMs are more difficult to initialize and parametrize (Smajgl et al., 2011); they are constantly criticized for lack of transparency and difficulty in evaluation (Müller et al., 2014); and they lead to difficulties in analysing and making sense of multidimensional output data (Lee et al., 2015). In light of this situation modellers are scrambling to develop standards for description like ODD and ODD+D (Grimm et al., 2006; Grimm et al., 2010; Müller et al., 2013) and deploy various tools to analyse output data (Lee et al., 2015). Meanwhile, scientific researchers are increasingly pondering the fundamental questions (O’Sullivan et al., 2015), such as: Is the model under development too simple or too complicated? How to define and quantify the level of complicatedness? What are the criteria for choosing simple vs. complicated models? What are the trade-offs between the two? Is there an appropriate or optimal level of complicatedness? If so, how to build an ABM with an appropriate level of complicatedness?

These questions, while relevant for any type of model, are particularly interesting and pressing for ABMs for several reasons. First, the agent-based modelling framework enables a considerable amount of flexibility (Bonabeau, 2002). Even simple ABMs are often more complicated compared to other well-established computational modelling approaches, such as system dynamics models. This is mainly due to the number and heterogeneity of entities represented the complexity of their individual behaviours and mutual interactions (Manson et al., 2012; O’Sullivan et al., 2012). In other words, developing ABMs entails a much larger number of model construction decisions than in other approaches, and modellers are thus more prone to building over-complicated models. Second, ABMs of complex coupled human-environmental systems need to represent various processes and data from both socio-economic and biophysical domains (Parker et al., 2008). Such models can easily reach a high level of complicatedness. Third, ABMs themselves are complex systems and the level of the complexity of a model can increase exponentially with the increasing details in model specifications. Ironically, ABMs are often criticized for simultaneously being too simple (regarding the rules and specifications) and too “complex” (mainly with respect to the model behaviours) (Conte and Paolucci, 2014). The acceptance of ABMs by a broader range of researchers and stakeholders has therefore been hindered, and some have cautioned that careful thought should be given to when ABM should be employed (O’Sullivan et al., 2012).

Furthermore, with the rapid improvements in data availability and computing power, modellers are increasingly building more and more complicated ABMs despite a general preference for “simple and nice” models. Modellers however rarely justify the level of complicatedness when building their ABMs, partially because of the lack of practical guidelines. There is an urgent need to clarify and discuss these issues (O’Sullivan et al., 2015; Wilensky and Rand, 2015). This paper aims to systematically elaborate on the still unresolved and frequently discussed issues of the appropriate level of complicatedness in model structure, focussing on ABMs of human-environmental systems. After examining the definition and quantification of model complicatedness, we discuss the trade-offs, problems and risks of complicated and simple models, and instigate a discussion on the appropriate level of complicatedness in ABMs. Finally, we offer some recommendations on best modelling practices in this regard.

## 2 Definition and quantification of model complicatedness

To avoid confusion and to form the basis for further discussion in the remainder of the paper, we first provide some clarification on terms and concepts and discuss the measurement of complicatedness.

### 2.1 Clarification of terminologies: complicatedness vs. complexity

The terms “complicated” and “complex” and the terms “complicatedness” and “complexity” are frequently used interchangeably in common usage. However, they have different connotations and are not synonyms in scientific and management contexts (Tang and Salminen, 2001). Complexity has even different meanings in different scientific fields, such as information theory, network theory, and software engineering. In the context of complex systems theory, complexity means the emergence and unpredictable behaviours exhibited by complex systems featured by many locally and often non-linearly interacting components. Model complexity can be either interpreted as structural complexity (as in Kolmogorov complexity (Kolmogorov, 1998)) or behavioural complexity (Casti, 1994). This can be confusing.

From the etymological perspective, “complicated” and “complex” share the same prefix “com”, which means together in its Proto-indo-european (PIE) root. But “complicated” comes from the Latin verb “plicare” which means fold (noun, as in ten folds) and implies many; “complex” has the Latin root, “plectere”, which means braiding, linking, weave, interlace, and intertwine. Thus, complex has the implication of interaction and intertwining besides the meaning of “many parts” shared with complicated.

By describing a model as complicated, we mean the model *structure* contains large numbers of variables or agents and, more importantly, detailed representation of processes and interactions via

logical rules and/or quantitative relationships (i.e. equations). Complicatedness accordingly refers to model entities and structures from the model construction perspective and is unrelated to cognitive and model behaviour aspects (Chwif et al., 2000). When referring to model structure, we advocate the usage of “complicated model” instead of “complex model”, as we do in the remainder of the paper.

In contrast, the terms complex and complexity largely refer to the model *behaviour* at the system level. Unlike complicatedness, complexity is not imposed by modellers but rather arises from interactions at the individual-level to produce simulated system-level behaviour that is emergent and possibly non-linear. Thus, the behaviour of an ABM is itself often complex and difficult to be described analytically. Thus, whereas the complicatedness of an ABM is a property of the model structure, the complexity of an ABM is a property of the model behaviour.

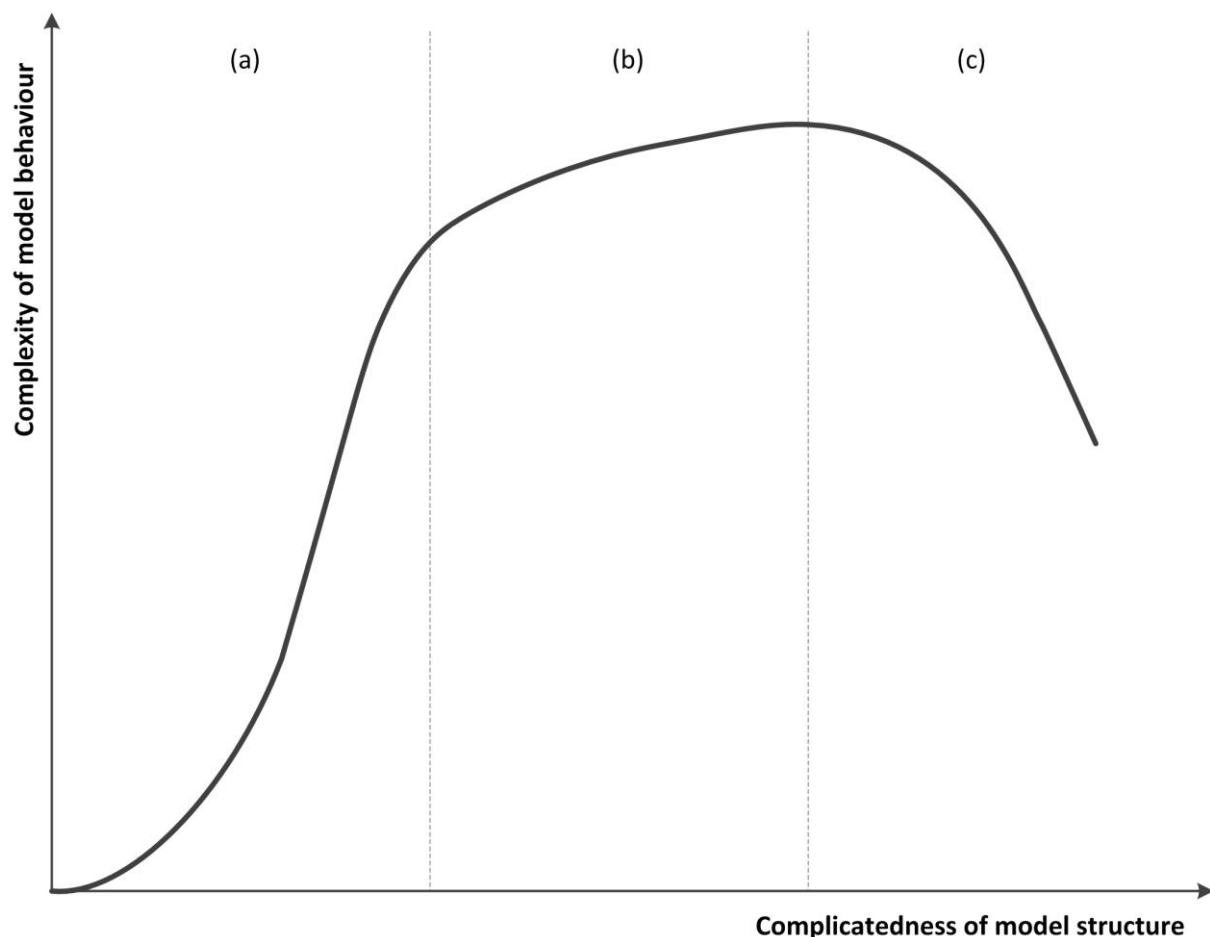


Fig. 1. Model complicatedness vs. Model complexity: (a) complexity increases exponentially with model complicatedness; (b) complexity increases at a lower ratio with model complicatedness; (c) complexity may decrease after certain threshold of model complicatedness

All ABMs should be complex in a sense that they should capture the key dynamics of real-world complex systems and therefore also show complex behaviour (Balbi and Giupponi, 2010). However,

from the model structure perspective ABMs can be simple or complicated. Of course, although the concepts “complicated” and “complex” differ, they are related (as shown in Fig. 1). For ABMs, simple models can still produce enormous complex behaviours as demonstrated in the classic “Game of Life” model (Conway, 1970), Schelling’s segregation model (Schelling, 1971), and CybErosion landform-evolution model (Wainwright, 2008). But the complexity of model behaviour may decrease after model complicatedness crosses a certain threshold. Reasons for such a negative relationship are possibly that the effect of too many entities cancels each other out, or that complex behaviour is blurred by many additive effects. In such cases, model behaviours tend to be less dynamic or even converge to equilibriums or regular patterns. Another reason is that complicated model specification may lead to chaotic or largely stochastic model behaviours. Both the ordered and chaotic patterns correspond to low level of complexity according to the concept of “edge of chaos” in the theory of complexity (Lewin, 1999). This seems counterintuitive relationship was confirmed by Hua and Pelikan (2012). After exploring several variations of the classic “Game of Life” model, they found the classic model, despite being simpler, surprisingly exhibits more diverse structure and more dynamic behaviours than some more complicated model variations. More in-depth researches however are needed to further investigate the relationship between the model complicatedness and complexity. The conceptual relationship illustrated in Fig. 1 is merely a hypothesis to be debated and tested.

## 2.2 Quantification of complicatedness

### 2.2.1 Factors influencing model complicatedness

The degree of complicatedness of an ABM depends on how detailed the model structure represents the modelled system and is determined by many factors. The sheer number and types of the entities (and their attributes), processes, and interactions all contribute to the complicatedness of an ABM. To summarize, we list the potential influencing factors based on the characterization of different model components specific to ABMs, largely following the ODD protocol (Grimm et al., 2010):

- **Agents’ representation:** Agents are the major entities of any ABM. The complicatedness of ABMs largely hinges on how many types of agents are present and how many attributes agents have. Simple agent-based land-use models (Millington et al., 2008; Müller et al., 2011; Parker and Meretsky, 2004; Wainwright, 2008) may have one type of agent (i.e., the land users/managers) described by relatively few state variables. Complicated models may have various types of agents (such as small-holders, state farms, cooperatives, policy makers, etc.), and even agent-groups, with heterogeneous characteristics (e.g., (Parker and Filatova, 2008), (Liu et al., 2006)). The number of state variables, which describe the properties of agents, might also be an instructive factor contributing to the complicatedness of the ABM (e.g., the

detailed socioeconomic attributes of farms modelled in AgriPoliS (Happe et al., 2006) and MP-MAS (Schreinemachers and Berger, 2011)).

- **Interactions:** Local and potentially complex interactions of agents with the environment is a defining feature of ABMs. The model may allow varying types of interactions among individuals (e.g., direct or indirect, unidirectional vs bidirectional, linear or non-linear, with or without feedback), which influences the overall complicatedness of ABMs.
- **Decision rules:** Decision rules of agents can vary from simple heuristics to optimization methods or detailed psychological models. Simple decision-making and behavioural rules can be represented with, for example, “if-then” rules or some simple mathematic equations. Complicated rules, on the other hand, may use sophisticated approaches such as Linear Programming (AgriPoliS (Happe et al., 2006); MP-MAS (Berger and Schreinemachers, 2006)), Decision Trees (LUDAS (Deadman et al., 2004)), multivariate regression (Schwarz and Ernst, 2009; Villamor et al., 2014) and Bayesian networks (IAMO-LUC (Sun and Müller, 2013)).
- **Environment:** The complicatedness of the environment is defined by its spatio-temporal resolution and the degree of detail at which the environment is characterized (e.g., land use types, structures, environmental processes). In many cases, a simple and artificial torus space is used as the spatial environment even in empirical studies without much justification. On the other hand, GIS layers are frequently used to provide a more realistic geographic settings for empirical ABMs (Crooks and Castle, 2012).
- **Scheduling:** The scheduling of a model defines the order of processes (Grimm et al., 2006). In the context of ABMs, modellers need to consider two update modes: synchronous vs. asynchronous. In the synchronous mode, all agents update their states simultaneously—their state changes are not seen by other agents until the next clock tick. Therefore, the order in which agents take actions does not matter. In asynchronous mode, conversely, agents update their states one by one; their updated state are immediately seen by other agents and may influence actions of other agents (Wilensky and Rand, 2015). In this case, the order of agents taking actions matters a lot to the model behaviour; models can produce rather different emergent patterns due to the introduction of delay factors (Caron-Lormier et al., 2008); more importantly, in this mode modellers also need to further specify whether agents act sequentially or in parallel. While asynchronous update seems more realistic, it is more complicated to specify and debug. .

It is worth noting that the above factor list is far from exhaustive—many more factors, such as the initialization process, agent behaviours besides decision-making, and formation of the sub-models, can all contribute to the complicatedness of ABMs. As ABMs are essentially one type of models or



modelling frameworks, it is difficult to consider all potential factors; the list above contains only the most prominent ones.

### 2.2.2 Measurements of model complicatedness

Although various ways of characterising model complicatedness exist, to our knowledge there are no well-accepted approaches, metrics or indices to quantify the complicatedness of an ABM. In algorithmic information theory model complicatedness can be approximated with the descriptive complexity or Kolmogorov complexity (Kolmogorov, 1998), which is basically a measurement of the resources needed to specify the model. In the case of ABMs, we propose to approximate the model complicatedness roughly by the length (e.g., number of lines or characters) of model code. Despite being quite rudimentary, researchers often use this natural, intuitive, and heuristic method. For example, when exploring an ABM in NetLogo®, one of the most popular ABM software platforms, researchers can browse through the model code to quickly estimate how complicated the model could be. Of course, this measurement is influenced by other factors, such as software choice or programming languages and quality/style of the coding. It works best when comparing models in the same platform or programming language. A standardized pseudo code should then be used to allow comparison of models implemented in different languages.

Another similar approach is to use model “size”, which accounts for model dimensionality, consecutiveness, and number of interacting processes, as an approximate measurement of the complicatedness of models (Boschetti, 2008). In the case of ABMs, the model size can be a function of the total number of agents, the number of agent types, the number of attributes of agents, the interactions among agents and between agents and environment, the number of sub-models describing agent behaviours, and so on. Even simpler, Nelles (2013) suggested that complicatedness of models is mainly related to number of parameters. Larger number of parameters implies a more complicated model. Clearly, if a model can be specifically formulated, for example, as a mathematic equation, the number of the parameters can be easily counted. As for ABMs, it is not that apparent given the heterogeneity and flexibility of ABMs. Here we propose the following general function:

$$MC = f(n_{parameters}) = f(gv + \sum_{i=1}^n sv_i + \sum_{j=1}^m ev_j) \quad (1)$$

where  $MC$  is the degree of model complicatedness;  $n_{parameters}$  is the total number of parameters whose values are allowed to varying between model executions;  $gv$  is the number of model-level variables;  $sv_i$  is the number of state or attribute variables of an agent type  $i$ ;  $n$  is the total number of all agent types;  $ev_j$  is the number of the parameters of a sub-model  $j$ ; and  $m$  is the total number of sub-models, which includes all agent behaviour models and other sub-models.

In this approach, parameters are considered equally: variable types and range of parameters are not accounted for; the formation of equations, linear or non-linear, is also ignored. Consequently, this approach, albeit more sophisticated than the above approach, also has limitations.

There are many other approaches worth exploring. For example, a potential approach is to represent models with an ontology graph (i.e., a graph showing all model entities as nodes and their relationships as edges) as described in Zhang et al. (2010), and then use the connectivity density, measured by the ratio of edges and nodes, as a proxy for complicatedness. However, few ABMs have accompanying ontologies, which limits the usefulness of this approach.

All these approaches have both advantages and limitations. The heterogeneity of ABMs makes the quantification of complicatedness and comparison of the degree of complicatedness among different ABMs extremely difficult. This is a great challenge as well as a pressing research area for the ABM modelling community. The proposed measurements here only scratch the surface of this topic, but it is beyond the scope of this paper to develop a consistent and comprehensive metric.

### 3 Systematic evaluation of simple and complicated ABMs

Simple and complicated models have different merits and weaknesses; building models at various levels of model complicatedness invariably involve trade-offs between generality, precision, and realism (Almaraz, 2014). The comparison of simple versus complicated ABMs reflects the general debate about the appropriate level of model detail needed to describe a specific system adequately. This debate is on-going in the scientific community (Brooks and Tobias, 1996; Evans et al., 2013; Jakeman et al., 2006; Janssen and Ostrom, 2006).

#### 3.1 General trade-offs between simple and complicated models

Human-environmental systems, such as land-use systems, are intrinsically complex. An increasing number of empirically-oriented agent-based models have been developed to better understand such complex systems. However, representing social interactions and environmental processes in excessive detail runs counter to the aim of understanding common trends in land change. The essence of modelling is to produce a simple representation of the real world so that it aids understanding (Carpenter, 2003). When an ABM is too complicated, however, the complexity of the model may approach that of the real world and the model itself thus becomes too difficult to understand and explore. This limits the usefulness of the model for improving system understanding. As a result, we lose the meaning of modelling (Couclelis, 2002; Peck, 2004).

Another potential risk for complicated ABMs is the problem of overfitting, which means that models are over specified and/or calibrated to a specific observation; models also explain random errors or

noises besides the actual underlying processes. As a result, models tend to perform extremely well on the training data, but fail miserably in other situations. Complicated ABMs, which have many parameters, tend to suffer from this problem and have therefore been criticized (Rand and Rust, 2011). Some complicated agent-based land-use change models, for example, may be able to reproduce historical land-use patterns in the study areas, but this alone cannot confirm the validity of the model. Simple models, in this regard, are less prone to the overfitting problem.

There are many more potential problems associated with complicated models and many scientists and modellers are critical towards complicated models.

On the other hand, simple models that follow the traditional principle of parsimony are not necessarily sufficient for understanding a complex system, as for instance in the case of land-use systems (Batty, 2007; O'Sullivan et al., 2015; O'Sullivan et al., 2012; Sohl and Claggett, 2013). Many detailed yet essential processes that influence land-use change processes, such as social networks (Manson et al., 2016) and land tenure rules, are difficult to represent in simulation models through simple structures. The social structures in which individuals are embedded vary widely across land use systems, are heavily context-dependent, and are not easily generalized (Rindfuss et al., 2004). Certainly, over-simplifying the context in which land-use decision-making is embedded can lead to incomplete and/or incorrect understanding of the forces that shape land-use choices (Magliocca et al., 2014). In addition, sticking to simple ABMs may also lead to a problem of under-exploiting the possibilities of ABMs (Conte and Paolucci, 2014).

In summary, ABMs of complex land use systems must strike a balance between generality and realism, and it necessitates trade-offs in model design.

### 3.2 Point-by-point evaluation of simple and complicated models

To facilitate discussion on the appropriate level of complicatedness of ABMs, we systematically compare simple and complicated ABMs from various perspectives summarized in three groups: (1) the purpose of the model; (2) the perspective of modellers, and (3) the perspective of model users. Some of these perspectives are independent from one another while others are closely related (e.g., prediction is, to a certain degree, a basis for decision making).

First, the **purpose of the model** is central and therefore guides the decision about the development of a simple or a more complicated ABM. Like any model, there are many possible alternative purposes for ABMs. Prediction, for example, is often commonly expected from users of simulation research, but multiple other reasons to model exist (Epstein, 2008; McBurney, 2012). Here we selected the most relevant purposes as follows:

- **Prediction:** Prediction implies quantitative forecasting under case-specific scenarios. This, however, often requires a highly-detailed representation of systems and their critical processes. Leaving out influential factors may distort predictive accuracy. Thus, predictive models may need to be complicated to achieve the required level of certainty. On the other hand, simple models, representing key processes, can be also used for predictive purposes, but rather qualitatively and not case-specific (Boero and Squazzoni, 2005). For example, Schelling's segregation model illustrates that highly segregated communities can result from relatively tolerant individual residents. But to predict the change of spatial patterns of a specific city will demand more detailed model specification.
- **Theory building:** Theory building aims to explain observations with new hypotheses or existing theories and to create, confirm or challenge existing theories (Epstein, 2008). ABMs can support theory building by combining inductive and deductive approaches in a so-called "third way of doing science" (Axelrod, 2006). To facilitate theory building with regard to complex systems, modelling usually involves generalization and simplification. Simple models elaborate on general questions ((Pace, 2003) – for ecosystem models) and are therefore more suitable to enhance theories that should possess general validity.
- **Decision-making:** ABMs may inform decision-makers by analysing policy scenarios and their consequences. Complicated models are suitable to support decision-making as they can be targeted at specific problems and developed for relevant stakeholders (Rounsevell et al., 2012). However, complicated models can be difficult to communicate, while simple models can be more favourable in cases where illustrative messages and narratives, as well as general rules about simple causal relationships, are more convincing to practitioners.
- **Case-specific analysis:** This is understood as the development of case-specific models using empirical data for a particular site, usually with the aim of quantitative prediction and scenario analysis. Simple models have difficulties representing site-specific characteristics of land systems with empirical data; more complicated ABMs are necessary and more suitable for such case-specific analysis than simple models (Peck, 2004).
- **Illuminate core dynamics:** To understand the dynamics of a system and reveal its critical underlying processes, simple models are often thought to be more suited. However, simple models may fail to deliver robust explanations under certain circumstances, at times raising the need for more complicated data-driven models to increase our understanding of system processes and the causalities between them.
- **Education:** ABMs have been widely used for educational purposes such as teaching students and informing real-world practitioners. By running the model, consequences of decisions may be explored without any significant risks. At first sight, simple models appear more

appropriate for educational purposes, considering the advantage of quick and easy model communication and transparency of model behaviour. However, this may change with the target group. Experts in the respective field might want to learn about the complex interplay of policy options, environmental scenarios, or behavioural assumptions, raising the need for more complicated models.

Second, another dimension in the discussion is the **perspective of modellers**, related to the specification, implementation, and analysis of a simulation model. More complicated simulation models are in general more challenging because they are difficult to comprehend and handle compared to simple models.

- **Conceptualization and implementation:** This comprises the development of the conceptual model design, data collection, and implementation as a computer program. Simple models usually demand less data which may facilitate their fast implementation. However, the simplification of processes must be well thought through which can be time-consuming. High diversity of entities and processes of complicated models demands more development time and resources. Furthermore, many complicated ABM are data-driven (O'Sullivan, 2008) and the data collection and parameterization needed often slows model implementation.
- **Calibration, verification, and validation:** After model implementation, model parameters may need to be adjusted to produce desirable patterns or values. This is the process of model calibration. Verification is the check for correctness of this implementation (that the model is built as intended), while validation checks the plausibility of model results (North and Macal, 2007). Simple models, with less parameters and simpler process representation, tend to be easier to calibrate and verify. In contrast, complicated models, with large parameter spaces and multiple interlinked sub-models and processes, can pose a daunting challenge for modellers to find sensible parameter values and ensure the model is implemented as designed. At the same time, complicated models are more prone to overfitting during the calibration process (as discussed above). With regard to validation, simple models are usually compared to existing theories or stylized patterns. The process is technically less challenging, but whether simple models really capture the key underlying processes is often difficult to tell. For complicated models, validation is a test of how a model reflects reality based on independently observed values and patterns. Considering the quantity of high-dimensional output data produced by complicated models, validation can be an arduous task.
- **Transferability:** Transferability includes whether the model can be applied to and has valid results across comparable contexts. The advantage of simple models regarding transferability

of the representation and derived results has been discussed previously (Parker et al., 2003; Rounsevell et al., 2012). This advantage is due to the fact that the questions addressed with the model are rather general, and thus the resulting patterns can be observed in different systems more generally. Evans et al. (2013), however, argue that only structurally realistic models (Grimm et al., 2004) can be flexible enough to describe characteristics of different study systems. In contrast, a simple model that fits one situation might not be able to account for differences in another situation and is therefore not transferable.

Finally, the **perspective of model users** is an important consideration for developing simple or complicated models. We identify two important constituents in simulation research: stakeholders and the scientific community. For successful simulation research, constituents need to trust the validity of the model and understand the results.

- **Transparency:** Transparency fosters the understanding of underlying processes and enables the interpretation of model results. Simple ABMs with less parameters, simpler rules and structures tend to be more transparent than complicated models. As a result, the model behaviour can be easily understood by peers and stakeholders. In turn, complicated models are often criticized as being “black boxes” (Topping et al., 2003) of the internal mechanics of which are unseen and therefore unknown.
- **Communication:** The communication of the model to other modellers and stakeholders is a major challenge of ABMs, because they may have diverse decision rules and methods that describe agents, and can be applied in a wide range of research domains (An, 2012; Bonabeau, 2002; Grimm et al., 2006; Müller et al., 2014). Complicated ABMs featuring various types of entities and processes tend to be more challenging to communicate and document than simple ABMs. Thus, assessment of complicated models is difficult, for example, during the peer-review publication process. Simple stylized (or “toy” models) are regarded as suitable to foster interdisciplinary communication, because joint model development forces participants from different disciplines to define concepts and make explicit their assumptions about relevant processes (Schlüter et al., 2013).
- **Trust/acceptance of stakeholders:** In order for stakeholders to build confidence in a model and its results, understanding model structure is a precondition. Otherwise, a practical impact of model findings is unlikely. On the one hand, stakeholders need to understand the model sufficiently (e.g., through transparent model design; see above), which demands simple models (Sohl and Claggett, 2013; Voinov and Bousquet, 2010). On the other hand, more complicated models ironically result from the request of stakeholders with various background and interests, such as questions like “Have you considered ...?”.

## 4 The appropriate level of complicatedness and implications for modelling

### 4.1 The principles of model building

As outlined above, ABMs can be structured to represent their target systems at a variety of levels of detail and complicatedness. What, then, is the “appropriate” level of complicatedness for ABMs? Minimal complicatedness of models is often advocated (as shown in Fig. 2 (A)). The principle of parsimony has been proposed in a variety of ways. These include Occam’s razor, where the simpler model is usually the preferred model if both simple and complicated models produce plausible explanations, and Einstein’s razor, which argues a model should be as simple as possible but not simpler (O’Sullivan et al., 2015). A vibrant example of this attitude is illustrated by the KISS principle (i.e., “keep it simple, stupid”) promoted by Axelrod (1997).

On the other hand, Edmonds and Moss (2004) challenge the KISS principle with a “Keep it Descriptive Stupid” (KIDS) approach. They argue that the model should be constructed complicated and detailed enough to model the richness of target systems; they also concede that mixtures of the KISS and KIDS approaches will be likely more appropriate. Similarly, O’Sullivan et al. (2015) have argued for a mid-level of complicatedness as the optimal or appropriate level. Grimm et al. (2005) used the term “Medawar zone” (originally appeared in Loehle (1990)), the intermediate range of complicatedness, to illustrate that the optimal level of complicatedness for ABMs falls somewhere in the middle (as shown in Fig. 2 (B)).

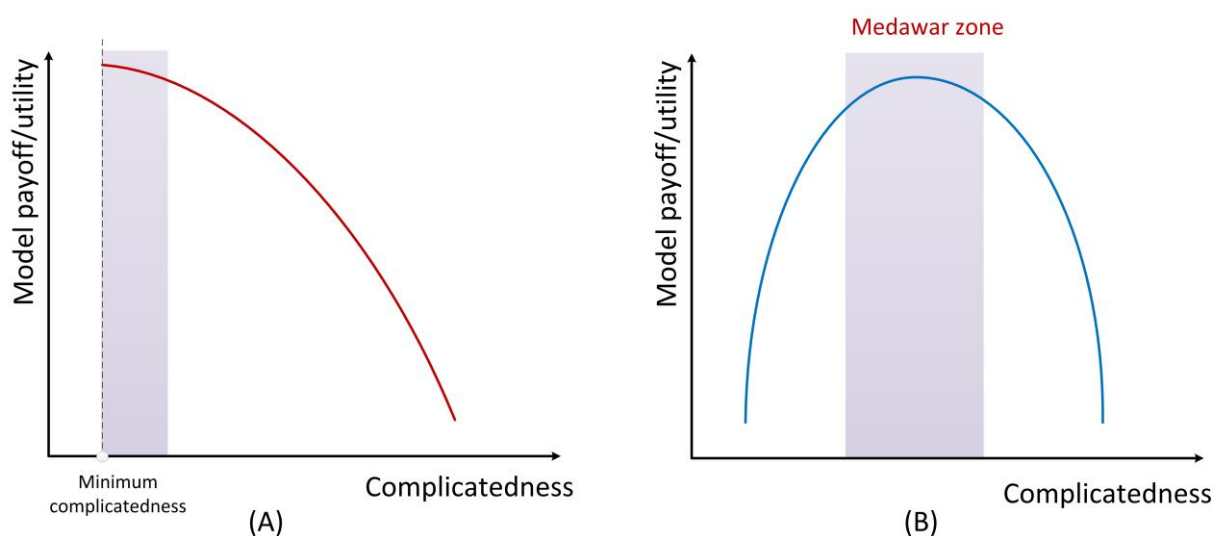


Fig. 2. Two classic modelling principles (figure (B) is modified from Grimm et al. (2005)). Shaded zones represent preferred complicatedness levels.

Conceptually, both the minimal simplicity and mid-level approaches, albeit seemingly contradicting one another, make great sense. There is no consensus on the appropriate level of complicatedness in the ABMs community. Correspondingly, there is no observed trend in movement towards simple or medium complicatedness models. Instead, increasing numbers of complicated empirically-grounded models have emerged to examine, for example, land-use change and environmental management (Janssen and Ostrom, 2006; Smajgl and Barreteau, 2014). At the same time, most newly developed ABMs, in particular in computational social science as characterised by articles in the *Journal of Artificial Societies and Social Simulation* (JASSS), are still on the simple end of the model spectrum (Conte and Paolucci, 2014).

We believe the conflicting view can be explained by the heterogeneity of ABMs—essentially, there are two types of ABMs, simple abstract models and empirically-grounded complicated models locating on both end of the model spectrum, with very different model purposes. To discuss the appropriate level of complicatedness, without clarifying the model types, and more precisely, model purposes, makes little sense and may result in false conclusions being drawn. Ultimately, the appropriate level of model complicatedness hinges on the model's intended purpose and the inherent nature of the system to be modelled. Depending on the specific research questions and how the model will be used, modellers can decide on whether a more complicated realistic model or a simple “fast-and frugal” model should be used (Carpenter, 2003; Van Nes and Scheffer, 2005).

Simple abstract models are generally intended for theory-building and explaining emergent patterns. They are often used to explore collective implications of individual-level decisions (e.g. Schelling (1969) for residential segregation; Brown and Robinson (2006) for urban land-use change)) to enrich our process-level understanding of a given system. In general, this type of model should follow the principle of parsimony and be constructed as simple as possible. Clear representation of important processes and interactions is a higher priority than the veracity of its outcomes.



On the other hand, empirically-grounded ABMs are generally more oriented towards prediction and often need to address specific questions posed by policy-makers at particular sites. This type of ABM needs to represent the detailed geographical settings and the processes therein, and are often data-intensive and driven by the site-specific research questions (e.g., agent-based land use models in Yucatán, Mexico by Manson (2006), Queensland, Australia by Valbuena et al. (2010), and Northern Ecuadorian Amazon Mena et al. (2011)). Consequently, these models cannot be too simple and must be sufficiently detailed and calibrated to the target land systems. Yet, modellers also need to be cautious not to build overly-complicated models, otherwise undue time and effort may be expended in data collection, model construction, testing (including comprehensive robustness and sensitivity analysis), and validation. Rather, modellers should try to build models which are as complicated as necessary to answer the specific research question---models in the “medawar zone”.

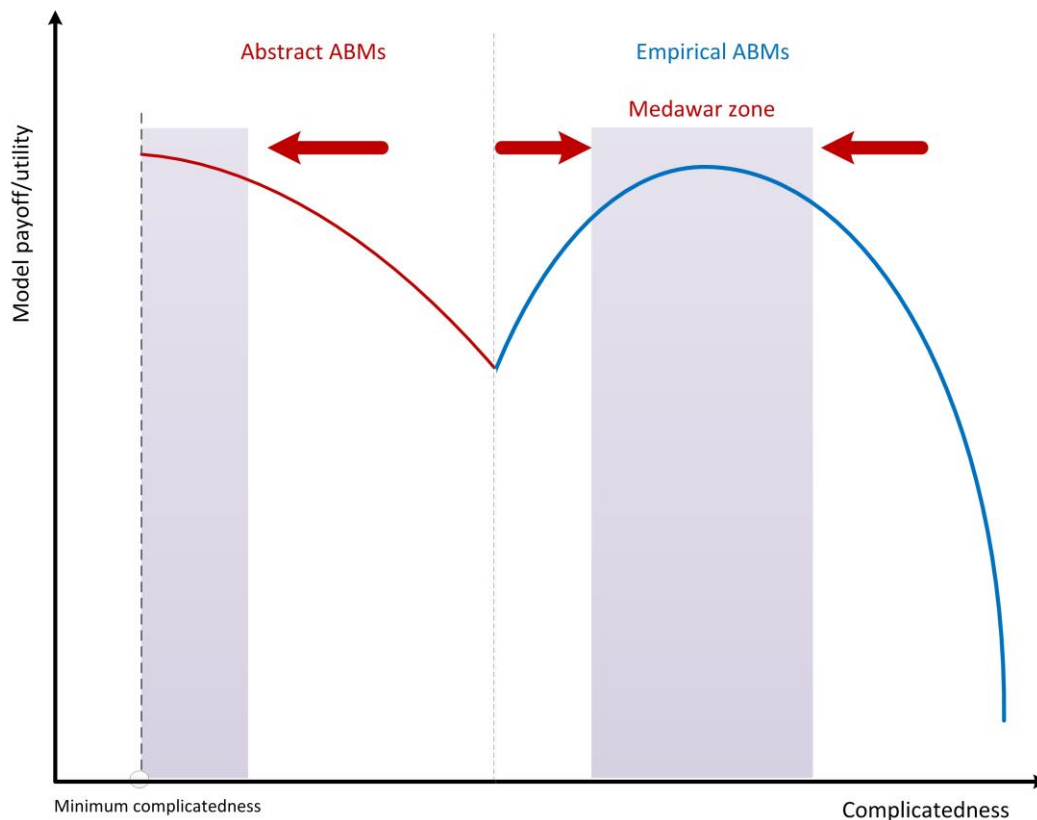


Fig. 3. The harmonization of principle of parsimony and “Medawar zone” in ABMs. The red line (left part of the curve) represents the utility for abstract models; the blue line (right part of the curve) represents the utility for empirical models.

We hence believe the two principles of minimum complicatedness and the “medawar zone” can both be applied in searching for the appropriate level of complicatedness, since they are relevant for a specific model type respectively. While the mid-level principle is more relevant for empirical ABMs

which need to be complicated to some degree, the principle of parsimony – whichever flavour – should be always followed (as we illustrate in Fig. 3).

Parsimony in model design and explanation of modelled phenomenon is desired for two reasons. First, land use systems are characterized by a multitude of interacting system components spanning multiple spatial and temporal scales, and thus our understanding of such systems is limited. The data required to parameterize the large number of free parameters associated with process-based models of complex systems is often unavailable or incomplete (Parker and Filatova, 2008). As Oreskes and Belitz (2001) argue, such a situation "opens the door to systematic error and bias" (p. 27). Thus, one objective of and justification for model parsimony is the minimization of the number of uncertain parameters and potential errors embedded in a model's design. Second, models must be constructed as parsimoniously as possible so that relationships between model inputs and outputs can be understood (Grimm et al., 2005; Parker and Filatova, 2008). After all, one of the main purposes of models is to elucidate important relationships and organizational structure of real systems that are often obscured by complexity (Greenberger et al., 1976). Simplifications must be made to represent real system process, or the model may be too complicated to interpret and useless for addressing the research question (Grimm et al., 2005; Parker et al., 2003).

#### 4.2 Practical issues influencing levels of complicatedness

The range of detail possible in ABMs means that all "levels of complicatedness" can seem *a priori* justifiable, and the question of whether any single level of representational detail for all models is "right" is ultimately moot. Agent-based modellers must use their prior knowledge of the system and their aims as a guide to the construction of their model, as there are no universal rules for defining the appropriate scope. While taking into account the general principles of model construction discussed above, there are practical considerations for establishing what the appropriate level of complicatedness is for any particular model:

- **Research questions:** The required degree of agent heterogeneity and their interactions, and hence the level of complicatedness of an ABM, should be dictated by the purpose of the model and given research questions. Ultimately, the appropriate level of representational detail of agents and their interactions is important for revealing or reproducing patterns of interest in the real world target system. For example, Valbuena et al. (2010) were interested in understanding how variation between types of farmer decision-making in response to global or regional processes influenced patterns of landscape structure and land use. This model may be assigned to the complicated end of the spectrum. In contrast, Millington (2012) set out to investigate how patterns of farmer decision-making might change as a result of cultural change due to constraints on agents' behaviour and their interactions. This

model, at the simple end of the model spectrum, represented different and fewer agent attributes and was less reliant on empirical data than that of Valbuena et al. (2010). Both choices on the level of model complicatedness, however, are appropriate given the differing modelling objectives. Whereas the complicated model aimed to examine policy impacts of agents' decision-making and consequences for land use change, the simple model was built to develop theory on how agent land use decision-making is related to behavioural change. Although both models represent similar real-world actors, the questions that motivated their use influenced their level of complicatedness.

- **Data availability and perspectives on theory vs. application:** The example above also highlights the importance of considering data availability and perspectives on theory vs. application. Valbuena et al. (2010) were primarily interested in using available data to help understand land use decisions in an applied sense. Data availability in the study region played an important role in the setup of the model structure. Millington (2012), in contrast, started from a theoretical question and used little data to inform model development. It is a challenge to parameterise models of individual decision-making at the level of individuals' internal psychological qualities, and ultimately it may not be possible to appropriately parameterise this level of representation with empirical data. However, the purpose of the model for exploring theory did not make such demands.
- **Computational constraints:** Beyond issues of data availability, resource considerations for implementation and analysis may play a (practical) role in deciding how to structure the model. Among these resource consideration, computational demands can pose problems for ABM modellers due to, for example, the large individual population and detailed representation of their behaviours (Bradhurst et al., 2016). For example, with the aim of being able to simulate land-use change over large extents (e.g., the entire European Union) the CRAFTY ABM framework needed to make seems unrealistic assumptions on agent behaviours to ensure computational efficiency was adequate for the purpose (Murray-Rust et al., 2014).

These examples demonstrate how the trade-off between objectives and resource availability may influence the appropriate level of detail for a given model to be established. Sometimes modellers may need to modify their desired level of complicatedness due to other non-scientific constraints and factors, such as the demands of stakeholders and the limitation of time, resources and funding.

### 4.3 Best modelling practices and strategies

The principles discussed above provide a valuable theoretical basis for modellers to identify the appropriate level of complicatedness for constructing their models. But these principles alone do not

provide a practical guide for which processes and variables should or should not be included in a model. As stated above, the appropriate level of model complicatedness is influenced by many practical factors, such as data availability, computational limitations, time, and funding resources. However, model purposes and research questions should always dictate how much detail and how complicated the ABM should be constructed. Therefore, clear and specific research questions and/or model purposes should be defined and framed as the starting point in the model design phase. The research question however does not translate to a specific level of complicatedness. Prior knowledge and experiences of the modeller may guide the choice of some model specifics such as the spatial and temporal resolution, scale, and boundaries of the model. But modellers need systemic approaches when facing substantial numbers of options and decisions as in the case of agent-based land use models. Here we list some useful modelling practices and strategies to help modellers to reach the appropriate level of complicatedness.

- **Pattern Oriented Modelling (POM):**

The identification and reproduction of characteristic patterns of the modelled system, through structural validation techniques like Pattern-Oriented Modelling (POM) (Grimm et al., 2005), can help guide which processes and variables must be included in model design to answer the driving research question. These characteristic patterns, for example, can be the temporal S-Curve during in adoption process, or the “fish-bone” spatial pattern in Amazon deforestation, or spatial clustering patterns measured by Moran’s I. These patterns will also inform the modeller about appropriate temporal and spatial scales. Ideally, target patterns identified at multiple hierarchical spatial and temporal scales are pulled from the empirical literature, and used to optimize model structure, test different theories of agent behaviour, and reduce parameter uncertainty (Grimm et al., 2005; Topping et al., 2010). Grimm et al. (2005) argue that “the key to understanding complex systems often lies in understanding how processes on different scales and hierarchical levels are bound to each other” (p. 988). Not only does the reproduction of simultaneous spatial and temporal patterns “raise the empirical bar” over model outcome validation techniques based on aggregate statistics (Brown et al., 2005), it can also help mitigate the influence of our personal interests, biases, and specific scales of perception in model design (Grimm et al., 2005) – ultimately leading to more parsimonious and hopefully insightful models.

- **Stepwise approach:** Agent-based land-use models often require socioeconomic and environmental components to depict complex human-environment interactions. Although these social, cognitive, and environmental processes may never be reliably generalized, simplified and encoded into a model, the effects of such processes can be tested indirectly

with an experimental approach. Starting with simple (prototype) models that capture processes that are simple and readily generalizable (e.g., environmental constraints on agriculture) can set a benchmark for the explanatory power of a relatively simple model. The simple prototype models can also allow for rapid hypothesis testing which can then be tested in more structurally rich models (cf. (Magliocca, 2015; Schlüter et al., 2013)). More detailed processes and model components can then be gradually added to the model structure such that the relative importance of each process can be quantified along the way. Eventually, a point is reached at which model performance fails to improve with additional processes and mechanisms (see also Buchmann et al. (2016) for the effect of various aspects being switched on or off on different patterns in a rather complicated ABM). Modellers can stop adding more details to the model.

On the other hand, in contrast to the common wisdom of starting simple, there are situations where the modeller starts from a more complicated model and removes processes and modules towards a simpler model. This approach can be vividly demonstrated with the abstraction process in Pablo Picasso artist creation of “Bull” (Fig. 4), where the details of a bull are gradually removed while maintaining the key structure. Starting with a complicated model can ensure critical components are not left out of the model (Buchmann et al., 2016). Furthermore, constructing a complicated model can paradoxically be easier than constructing a well-contested simple one. Starting complicated might be easier, for example, if code for model modules already exists or if characteristics of the target system cannot initially be well reproduced by a highly simplified model.

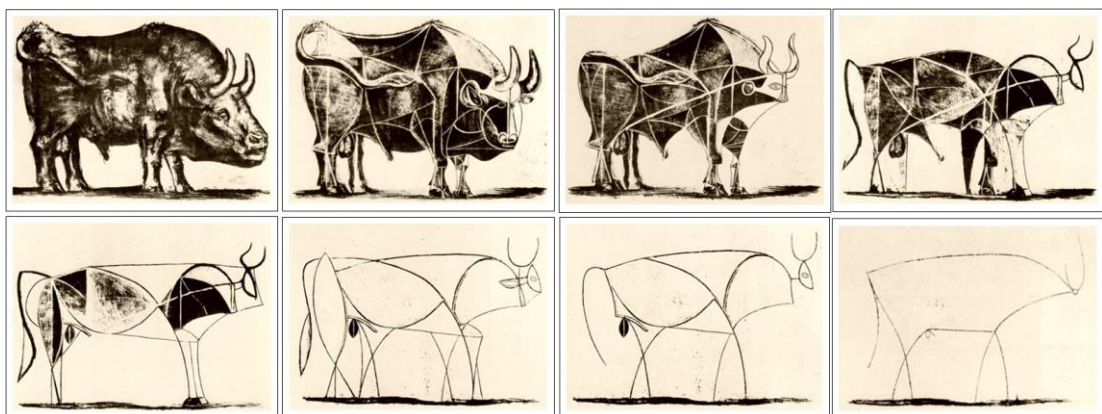


Fig. 4. The simplification of a “Bull” by Pablo Picasso, Museum of Modern Art, New York, US

(8 out of 11 lithographs are selected to illustrate the simplification process; images are downloaded from artfactory.com with written permissions; copyright: succession Picasso / VG Bild-Kunst, Bonn 2016)

During the stepwise process, statistical models and sensitivity analysis can be used to evaluate whether a factor, a process or module should be included or excluded. For example, Ligmann-Zielinska et al. (2014) demonstrated that simpler land-use models can be produced while retaining the performance of the model with the sensitivity analysis in their empirical research in Michigan, USA. Modellers should have in mind though that generic stepwise model selection based on statistical benchmarks might have some unintended results. For example, some well-known causal relations can be removed, which may cause a loss of generality and transferability of the model. The domain knowledge should be used to complement the rigid statistical process.

- **Parsimony during all stages of modelling cycle:** ABM modellers need to ensure a habit of parsimony and carefully design and construct the models to avoid any unnecessary details throughout the planning, implementation, and testing phases. The control of complicatedness in the design and planning phases is in particular critical because resources and time can be saved for later stages, and the reduction of complicatedness in an implemented model is time-consuming and difficult. Therefore, an extensive planning phase is important to reduce the need for later adaptations. While it is easier said than done, there are useful tools to facilitate the simplification process in different phases.

First, the modelling process is largely akin to software development: representing complicated target systems while trying to avoid unnecessary complicatedness, time-wasting, errors, and artefacts. Thus, software engineering methods and approaches, which are applied and approved in many ways, may serve as useful orientation to improve the efficiency and validity of ABMs with potentially high complicatedness. During model design, for example, the use of diagrams is especially helpful to identify most relevant and critical parts in the model. Modellers can then think about possible adaptations and simplifications of the model. We advocate using UML (Unified Model Language) diagrams, which provide a variety of standard schemes and visual illustrations for planning and communicating processes, and causal loop diagrams, which help to understand the causal relationships and identify the critical processes.

A second useful approach from software engineering is a modular design approach that subdivides a complicated system, in our case, ABMs, into independent sub-modules. A modular program structure that was planned beforehand provides some useful features, i.e., the use of independent submodules which can be switched on and off. Also, existing tested and verified sub-routines and modules from other models or well-known model-libraries can

be reused in different ABMs (Bell et al., 2015; Boulaire et al., 2015). Such modular approaches can facilitate the aforementioned stepwise approach. This search for modularity might also affect the level of interoperability among different models and platforms, currently negligible in ABM research. For a successful modularization it is important to define interfaces between the submodules, so that they can be integrated easily. Possibly, it might be beneficial to use a framework such as Repast Symphony or NetLogo with abundant coding samples and a model library full of simple models.

## 5 Discussion and conclusions

We have argued that complexity and complicatedness are different concepts in agent-based modelling. The model complicatedness refers to the detailedness of model structure; the model complexity is to describe the behaviours of an ABM. Although the capacity of capturing complex dynamics is a desirable attribute of all ABMs, complicatedness pertains to model construction and depends on the level of detail in the representation of the target system. Existing ABMs of human-environmental systems vary between simple and complicated models (Matthews et al., 2007; O'Sullivan et al., 2015; O'Sullivan et al., 2012). Despite the fact that the simple structure of ABMs and the complex behaviours they could reproduce drove their popularity in the first place, now there is a clear shift towards empirically-grounded ABMs enticed by the increasing availability of data, computational power, and various requests from stakeholders for policy and scenario analysis. This increasing complicatedness of ABMs comes with potential risks and entails trade-offs. We systematically assessed the merits and demerits of simple and complicated, empirically-grounded models from various perspectives.

We suggest that simple models are preferred for theory-building and education purposes as they present advantages in terms of transparency and prevention of overfitting. Complicated models seem inevitable given the complexity of the targeted system (e.g., land-use systems) especially when used for case-specific analysis. They are also more likely to be useful in terms of prediction capacity, scenario analysis and decision making support. On the contrary, simple models tend to be more accessible in terms of implementation and testing, transferability, and communication.

We firmly believe ABMs should be constructed as simple as possible, and as complicated as necessary to address clearly predefined and described research questions. The POM approach can help modellers to decide how much detail is enough by using patterns found in the real systems to frame the research questions. Thus, the POM approach can likely lead to an appropriate level of complicatedness. In addition, unnecessary complicatedness can be avoided by carefully planning



model development and communication (Jonker and Treur, 2013). Although it is still challenging to define and reach an appropriate level of complicatedness for ABMs, we nevertheless present modellers with some other useful strategies and approaches, such as stepwise and modular approaches.

As for the future direction of agent-based modelling of human-environmental systems, we believe modellers should focus on both ends of the model complicatedness spectrum where the payoff and usefulness of the models are higher. The “mid-level” models, which try to compromise between simple and complicated models by, for example, combining empirical data with stylized behavioural rules, may also compromise the objective of ABMs. The empirical details in such models may be unnecessary and even hinder the theory building purpose; on the other hand, the stylized components may cripple the prediction capability and then limit their applications in policy support. Therefore, modellers need to develop simple theory-oriented models, still with rich dynamics in terms of model behaviours, for understanding the key processes of land systems. Surprisingly, we struggled to find such simple and elegant agent-based land-use models in the latest literature (Magliocca et al., 2013; Magliocca and Ellis, 2013). The key challenge might be to find the interesting research questions and the theories to test and explore. Meanwhile, modellers should not be intimidated by the risks involved in developing complicated models, and empirically-grounded ABMs can provide useful insights for stakeholders. The capacity of ABMs should not be wasted (Conte and Paolucci, 2014). The increasing development of more powerful tools such as statistical methods, meta modelling, POM, visualization, artificial intelligence, and strict protocols like ODD can help modellers in dealing with the complicatedness in the model construction and output analysis (Lee et al., 2015). However, if complicated models stay purely site-specific and data-driven, the scientific contribution of such models will be limited. Modellers should aim to ground their models on solid theories and, more importantly, endeavour to generalize new theories from complicated simulation models. Although previously recognised (Perry and Bond, 2004), this remains an ongoing challenge. A practical aim should be the development of a common metric to measure the complicatedness of ABMs quantitatively so that different models can be compared. Although our discussions are focused on agent-based land use models, many valuable lessons have been drawn from various modelling fields. We are confident that the insights and guidelines provided in this article, albeit open to ongoing debate, are also applicable and helpful to agent-based modelling of different research domains, and even other simulation approaches in general.

## Acknowledgement



We are grateful to the valuable comments and suggestions from five anonymous reviewers. We thank all participants (including Gary Polhill, Dawn Parker, Arika Ligmann-Zielinska, Evgeny Latynskiy, etc.) of our workshop during iEMSs 2014 for the insightful discussions and input. ZS was supported by the European Commission (Collaborative Project Impacts of Reducing Emissions from Deforestation and Forest Degradation and Enhancing Carbon Stocks, I-REDD+, FP7-ENV-2010-265286). NM was supported by the National Socio-Environmental Synthesis Center (SESYNC) through a grant from the U.S. National Science Foundation (# DBI-1052875) to the University of Maryland. CB acknowledges financial support by the German Ministry for Education and Research (BMBF) for the project “Gemeinsam auf dem Weg in die energieeffiziente urbane Moderne – Einrichtung eines akteursorientierten Energiemanagementsystems in Delitzsch” (ref. no. 03SF0408A). BM was supported by the German Federal Ministry of Education and Research (BMBF—01LN1315A) within the Junior Research Group POLISES.

## 6 References

- Albin, P., Foley, D.K., 1992. Decentralized, dispersed exchange without an auctioneer: A simulation study. *Journal of Economic Behavior & Organization* 18(1) 27-51.
- Almaraz, P., 2014. Simple Models, Complex Models, Useful Models: Can We Tell Them from the Flap of a Butterfly's Wings? *Frontiers in Ecology and Evolution* 2.
- An, L., 2012. Modeling human decisions in coupled human and natural systems: review of agent-based models. *Ecological Modelling* 229 25-36.
- Axelrod, R., 2006. *Simulation in the social sciences*.
- Balbi, S., Giupponi, C., 2010. Agent-based modelling of socio-ecosystems: a methodology for the analysis of adaptation to climate change. *International Journal of Agent Technologies and Systems* 2(4) 17-38.
- Balbi, S., Giupponi, C., Perez, P., Alberti, M., 2013. A spatial agent-based model for assessing strategies of adaptation to climate and tourism demand changes in an alpine tourism destination. *Environmental Modelling & Software* 45 29-51.
- Batty, M., 2007. *Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals*. The MIT press.
- Bell, A.R., Robinson, D.T., Malik, A., Dewal, S., 2015. Modular ABM development for improved dissemination and training. *Environmental Modelling & Software* 73 189-200.

- Berger, T., Schreinemachers, P., 2006. Creating agents and landscapes for multiagent systems from random samples. *Ecology and Society* 11(2) 19.
- Boero, R., Squazzoni, F., 2005. Does empirical embeddedness matter? Methodological issues on agent-based models for analytical social science. *Journal of Artificial Societies and Social Simulation* 8(4).
- Bonabeau, E., 2002. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences* 99(suppl 3) 7280-7287.
- Boschetti, F., 2008. Mapping the complexity of ecological models. *Ecological Complexity* 5(1) 37-47.
- Boulaire, F., Utting, M., Drogemuller, R., 2015. Dynamic agent composition for large-scale agent-based models. *Complex Adaptive Systems Modeling* 3(1) 1-23.
- Bradhurst, R.A., Roche, S.E., East, I.J., Kwan, P., Garner, M.G., 2016. Improving the computational efficiency of an agent-based spatiotemporal model of livestock disease spread and control. *Environmental Modelling & Software* 77 1-12.
- Brooks, R.J., Tobias, A.M., 1996. Choosing the best model: Level of detail, complexity, and model performance. *Mathematical and computer modelling* 24(4) 1-14.
- Brown, D.G., Page, S., Riolo, R., Zellner, M., Rand, W., 2005. Path dependence and the validation of agent - based spatial models of land use. *International Journal of Geographical Information Science* 19(2) 153-174.
- Brown, D.G., Robinson, D.T., 2006. Effects of heterogeneity in residential preferences on an agent-based model of urban sprawl. *Ecology and Society* 11(1) 46.
- Buchmann, C.M., Grossmann, K., Schwarz, N., 2016. How agent heterogeneity, model structure and input data determine the performance of an empirical ABM – A real-world case study on residential mobility. *Environmental Modelling & Software* 75 77-93.
- Caron-Lormier, G., Humphry, R.W., Bohan, D.A., Hawes, C., Thorbek, P., 2008. Asynchronous and synchronous updating in individual-based models. *Ecological Modelling* 212(3) 522-527.
- Carpenter, S.R., 2003. The need for fast-and-frugal models. *Models in Ecosystem Science*, CD Canham, JJ Cole and WK Laurenroth (eds.), Princeton University Press, Princeton, New Jersey 455-460.
- Casti, J.L., 1994. Simple and complex models in science.
- Chwif, L., Barretto, M.R.P., Paul, R.J., 2000. On simulation model complexity, *Proceedings of the 32nd conference on Winter simulation*. Society for Computer Simulation International, pp. 449-455.
- Conte, R., Paolucci, M., 2014. On agent-based modeling and computational social science. *Frontiers in Psychology* 5 668.
- Conway, J., 1970. The game of life. *Scientific American* 223(4) 4.

- Couclelis, H., 2002. Why I no longer work with Agents, In: Parker, D.C., Berger, T., Manson, S.M. (Eds.), *Agent-based models of land-use and land-cover change: Report and review of an international workshop*, Irvine, Oct 4-7, 2001 LUCC International Project Office: Louvain-la-neuve, Belgium, pp. 3-5.
- Crooks, A., Castle, C., Batty, M., 2008. Key challenges in agent-based modelling for geo-spatial simulation. *Computers, Environment and Urban Systems* 32(6) 417-430.
- Crooks, A.T., Castle, C.J., 2012. The integration of agent-based modelling and geographical information for geospatial simulation, *Agent-based models of geographical systems*. Springer, pp. 219-251.
- Deadman, P., Robinson, D., Moran, E., Brondizio, E., 2004. Colonist household decisionmaking and land-use change in the Amazon Rainforest: an agent-based simulation. *Environment and Planning B* 31 693-710.
- Edmonds, B., Moss, S., 2004. From KISS to KIDS—an ‘anti-simplistic’ modelling approach, *International Workshop on Multi-Agent Systems and Agent-Based Simulation*. Springer, pp. 130-144.
- Epstein, J.M., 2008. Why Model? *Journal of Artificial Societies and Social Simulation* 11(4) 12.
- Epstein, J.M., Axtell, R., 1996. *Growing artificial societies: social science from the bottom up*. Brookings Institution Press.
- Evans, M.R., Grimm, V., Johst, K., Knuuttila, T., de Langhe, R., Lessells, C.M., Merz, M., O'Malley, M.A., Orzack, S.H., Weisberg, M., 2013. Do simple models lead to generality in ecology? *Trends in ecology & evolution* 28(10) 578-583.
- Filatova, T., Verburg, P.H., Parker, D.C., Stannard, C.A., 2013. Spatial agent-based models for socio-ecological systems: Challenges and prospects. *Environmental Modelling & Software* 45 1-7.
- Greenberger, M., Crenson, M.A., Crissey, B.L., 1976. *Models in the policy process: Public decision making in the computer era*. Russell Sage Foundation:[distributed by Basic Books].
- Grimm, V., 1999. Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future? *Ecological Modelling* 115(2-3) 129-148.
- Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S.K., Huse, G., Huth, A., Jepsen, J.U., Jørgensen, C., Mooij, W.M., Müller, B., Pe'er, G., Piou, C., Railsback, S.F., Robbins, A.M., Robbins, M.M., Rossmanith, E., Rüger, N., Strand, E., Souissi, S., Stillman, R.A., Vabø, R., Visser, U., DeAngelis, D.L., 2006. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling* 198(1-2) 115-126.
- Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J., Railsback, S.F., 2010. The ODD protocol: A review and first update. *Ecological Modelling* 221(23) 2760-2768.
- Grimm, V., Lorek, H., Finke, J., Koester, F., Malachinski, M., Sonnenschein, M., Moilanen, A., Storch, I., Singer, A., Wissel, C., 2004. META-X: generic software for metapopulation viability analysis. *Biodiversity & Conservation* 13(1) 165-188.

- Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S.F., Thulke, H.-H., Weiner, J., Wiegand, T., DeAngelis, D.L., 2005. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *science* 310(5750) 987-991.
- Happe, K., Kellermann, K., Balmann, A., 2006. Agent-based analysis of agricultural policies: an illustration of the Agricultural Policy Simulator AgriPoliS, its adaptation and behavior. *Ecology & Society* 11(1) 49.
- Hua, D., Pelikan, M., 2012. 'Variations on Conway's Game of Life and Other Cellular Automata'.
- Jakeman, A.J., Letcher, R.A., Norton, J.P., 2006. Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling & Software* 21(5) 602-614.
- Janssen, M.A., Ostrom, E., 2006. Empirically based, agent-based models. *Ecology and Society* 11(2).
- Jonker, C.M., Treur, J., 2013. A formal approach to building compositional agent-based simulations, *Simulating Social Complexity*. Springer, pp. 57-94.
- Kolmogorov, A.N., 1998. On tables of random numbers. *Theoretical Computer Science* 207(2) 387-395.
- Lee, J.-S., Filatova, T., Ligmann-Zielinska, A., Hassani-Mahmooei, B., Stonedahl, F., Lorscheid, I., Voinov, A., Polhill, G., Sun, Z., Parker, D.C., 2015. The Complexities of Agent-Based Modeling Output Analysis. *Journal of Artificial Societies and Social Simulation* 18(4) 4.
- Lewin, R., 1999. *Complexity: Life at the edge of chaos*. University of Chicago Press.
- Ligmann-Zielinska, A., Kramer, D.B., Spence Cheruvilil, K., Soranno, P.A., 2014. Using Uncertainty and Sensitivity Analyses in Socioecological Agent-Based Models to Improve Their Analytical Performance and Policy Relevance. *PLoS ONE* 9(10) e109779.
- Liu, X., Li, X., Anthony, G.-O., 2006. Multi-agent systems for simulating spatial decision behaviors and land-use dynamics. *Science in China Series D: Earth Sciences* 49(11) 1184-1194.
- Loehle, C., 1990. A guide to increased creativity in research: inspiration or perspiration? *Bioscience* 123-129.
- Magliocca, N.R., 2015. Model-Based Synthesis of Locally Contingent Responses to Global Market Signals. *Land* 4(3) 807-841.
- Magliocca, N.R., Brown, D.G., Ellis, E.C., 2013. Exploring agricultural livelihood transitions with an agent-based virtual laboratory: global forces to local decision-making. *PLoS ONE* 8(9) e73241.
- Magliocca, N.R., Brown, D.G., Ellis, E.C., 2014. Cross-Site Comparison of Land-Use Decision-Making and Its Consequences across Land Systems with a Generalized Agent-Based Model. *PLoS ONE* 9(1) e86179.
- Magliocca, N.R., Ellis, E.C., 2013. Using Pattern - oriented Modeling (POM) to Cope with Uncertainty in Multi - scale Agent - based Models of Land Change. *Transactions in GIS* 17(6) 883-900.

- Manson, S., 2006. Land use in the southern Yucatán peninsular region of Mexico: Scenarios of population and institutional change. *Computers, Environment and Urban Systems* 30(3) 230-253.
- Manson, S.M., Jordan, N.R., Nelson, K.C., Brummel, R.F., 2016. Modeling the effect of social networks on adoption of multifunctional agriculture. *Environmental Modelling & Software* 75 388-401.
- Manson, S.M., Sun, S., Bonsal, D., 2012. Agent-based modeling and complexity, Agent-based models of geographical systems. Springer, pp. 125-139.
- Marks, R.E., 1992. Breeding hybrid strategies: Optimal behaviour for oligopolists. *Journal of Evolutionary Economics* 2(1) 17-38.
- Matthews, R., Gilbert, N., Roach, A., Polhill, J., Gotts, N., 2007. Agent-based land-use models: a review of applications. *Landscape Ecology* 22(10) 1447-1459.
- McBurney, P., 2012. What Are Models for?, In: Cossentino, M., Kaisers, M., Tuyls, K., Weiss, G. (Eds.), *Multi-Agent Systems*. Springer Berlin Heidelberg, pp. 175-188.
- Mena, C.F., Walsh, S.J., Frizzelle, B.G., Xiaozheng, Y., Malanson, G.P., 2011. Land use change on household farms in the Ecuadorian Amazon: Design and implementation of an agent-based model. *Applied Geography* 31(1) 210-222.
- Millington, J., Romero-Calcerrada, R., Wainwright, J., Perry, G., 2008. An agent-based model of Mediterranean agricultural land-use/cover change for examining wildfire risk. *Journal of Artificial Societies and Social Simulation* 11(4) 4.
- Millington, J.D.A., 2012. Using social psychology theory for modelling farmer decision-making, In: Seppelt, R., Voinov, A., Lange, S. (Eds.), *International Environmental Modelling and Software Society (iEMSs) 2012 International Congress on Environmental Modelling and Software. Managing Resources of a Limited Planet: Pathways and Visions under Uncertainty, Sixth Biennial Meeting: Leipzig, Germany*, pp. 2485-2492.
- Müller, B., Balbi, S., Buchmann, C.M., De Sousa, L., Dressler, G., Groeneveld, J., Klassert, C.J., Le, Q.B., Millington, J.D., Nolzen, H., 2014. Standardised and transparent model descriptions for agent-based models: Current status and prospects. *Environmental Modelling & Software* 55 156-163.
- Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., Schlüter, M., Schulze, J., Weise, H., Schwarz, N., 2013. Describing human decisions in agent-based models—ODD+ D, an extension of the ODD protocol. *Environmental Modelling & Software* 48 37-48.
- Müller, B., Quaas, M.F., Frank, K., Baumgärtner, S., 2011. Pitfalls and potential of institutional change: Rain-index insurance and the sustainability of rangeland management. *Ecological Economics* 70(11) 2137-2144.
- Murray-Rust, D., Brown, C., van Vliet, J., Alam, S.J., Robinson, D.T., Verburg, P.H., Rounsevell, M., 2014. Combining agent functional types, capitals and services to model land use dynamics. *Environmental Modelling & Software* 59 187-201.

- Nelles, O., 2013. Nonlinear system identification: from classical approaches to neural networks and fuzzy models. Springer Science & Business Media.
- North, M.J., Macal, C.M., 2007. Managing business complexity: discovering strategic solutions with agent-based modeling and simulation. Oxford University Press.
- O'Sullivan, D., 2008. Geographical information science: agent-based models. Progress in Human Geography.
- O'Sullivan, D., Evans, T., Manson, S., Metcalf, S., Ligmann-Zielinska, A., Bone, C., 2015. Strategic directions for agent-based modeling: avoiding the YAAWN syndrome. Journal of Land Use Science(ahead-of-print) 1-11.
- O'Sullivan, D., Millington, J., Perry, G., Wainwright, J., 2012. Agent-Based Models—Because They're Worth It?, Agent-based models of geographical systems. Springer, pp. 109-123.
- Oreskes, N., Belitz, K., 2001. Philosophical issues in model assessment, In: Anderson, M.G., Bates, P.D. (Eds.), Model validation: Perspectives in hydrological science. John Wiley and Sons: London, pp. 23-41.
- Pace, M.L., 2003. The utility of simple models in ecosystem science. Models in Ecosystem Science. Princeton University Press, Princeton, NJ 49-62.
- Parker, D.C., Filatova, T., 2008. A conceptual design for a bilateral agent-based land market with heterogeneous economic agents. Computers, Environment and Urban Systems 32(6) 454-463.
- Parker, D.C., Hessel, A., Davis, S.C., 2008. Complexity, land-use modeling, and the human dimension: Fundamental challenges for mapping unknown outcome spaces. Geoforum 39(2) 789-804.
- Parker, D.C., Manson, S.M., Janssen, M.A., Hoffmann, M.J., Deadman, P., 2003. Multi-agent systems for the simulation of land-use and land-cover change: A review. Annals of the Association of American Geographers 93(2) 314-337.
- Parker, D.C., Meretsky, V., 2004. Measuring pattern outcomes in an agent-based model of edge-effect externalities using spatial metrics. Agriculture, Ecosystems & Environment 101(2) 233-250.
- Peck, S.L., 2004. Simulation as experiment: a philosophical reassessment for biological modeling. Trends in ecology & evolution 19(10) 530-534.
- Perry, G.L., Bond, N.R., 2004. Spatial population models for animals, In: Wainwright, J., Mulligan, M. (Eds.), Environmental modelling: finding simplicity in complexity. Wiley, pp. 157-169.
- Pierr, A., Ungaro, F., Ciancaglini, A., Happe, K., Sahrbacher, A., Sattler, C., Uthes, S., Zander, P., 2009. Integrated assessment of future CAP policies: land use changes, spatial patterns and targeting. Environmental Science & Policy 12(8) 1122-1136.
- Poppy, G., Chiotha, S., Eigenbrod, F., Harvey, C., Honzák, M., Hudson, M.D., Jarvis, A., Madise, N., Schreckenberg, K., Shackleton, C., 2014. Food security in a perfect storm: using the ecosystem

- services framework to increase understanding. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369(1639) 20120288.
- Rand, W., Rust, R.T., 2011. Agent-based modeling in marketing: Guidelines for rigor. *International Journal of Research in Marketing* 28(3) 181-193.
- Reynolds, C.W., 1987. Flocks, herds and schools: A distributed behavioral model. *ACM Siggraph Computer Graphics* 21(4) 25-34.
- Richards, R., Sahinb, O., Sanoa, M., Meyneckea, J.-O., Tillere, R., 2014. App2Adapt: Using Tablet Technology to Elicit Conditional Probabilities for Bayesian Belief Network Modelling, In: Ames, D.P., Quinn, N.W.T., Rizzoli, A.E. (Eds.), *The 7th International Congress on Environmental Modelling and Software*: San Diego, California, USA.
- Rindfuss, R.R., Walsh, S.J., Turner, I.B.L., Fox, J., Mishra, V., 2004. Developing a science of land change: Challenges and methodological issues. *Proceedings of the National Academy of Sciences* 101(39) 13976-13981.
- Rounsevell, M.D.A., Robinson, D.T., Murray-Rust, D., 2012. From actors to agents in socio-ecological systems models. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367(1586) 259-269.
- Schelling, T.C., 1969. Models of segregation. *The American Economic Review* 488-493.
- Schelling, T.C., 1971. Dynamic models of segregation†. *Journal of mathematical sociology* 1(2) 143-186.
- Schlüter, M., Müller, B., Frank, K., 2013. How to use models to improve analysis and governance of social-ecological systems-the reference frame MORE. Available at SSRN 2037723.
- Schreinemachers, P., Berger, T., 2011. An agent-based simulation model of human–environment interactions in agricultural systems. *Environmental Modelling & Software* 26(7) 845-859.
- Schwarz, N., Ernst, A., 2009. Agent-based modeling of the diffusion of environmental innovations—an empirical approach. *Technological forecasting and social change* 76(4) 497-511.
- Smajgl, A., Barreteau, O., 2014. Empiricism and Agent-Based Modelling, *Empirical Agent-Based Modelling-Challenges and Solutions*. Springer, pp. 1-26.
- Smajgl, A., Brown, D.G., Valbuena, D., Huigen, M.G.A., 2011. Empirical characterisation of agent behaviours in socio-ecological systems. *Environmental Modelling & Software* 26(7) 837-844.
- Sohl, T.L., Claggett, P.R., 2013. Clarity versus complexity: Land-use modeling as a practical tool for decision-makers. *Journal of environmental management* 129 235-243.
- Sun, Z., Müller, D., 2013. A framework for modeling payments for ecosystem services with agent-based models, Bayesian belief networks and opinion dynamics models. *Environmental Modelling & Software* 45(0) 15-28.

- Tang, V., Salminen, V., 2001. Towards a theory of complicatedness: framework for complex systems analysis and design.
- Topping, C.J., Hansen, T.S., Jensen, T.S., Jepsen, J.U., Nikolajsen, F., Odderskær, P., 2003. ALMaSS, an agent-based model for animals in temperate European landscapes. *Ecological Modelling* 167(1) 65-82.
- Topping, C.J., Høye, T.T., Olesen, C.R., 2010. Opening the black box—Development, testing and documentation of a mechanistically rich agent-based model. *Ecological Modelling* 221(2) 245-255.
- Valbuena, D., Verburg, P.H., Bregt, A.K., Ligtenberg, A., 2010. An agent-based approach to model land-use change at a regional scale. *Landscape Ecology* 25(2) 185-199.
- Van Nes, E.H., Scheffer, M., 2005. A strategy to improve the contribution of complex simulation models to ecological theory. *Ecological Modelling* 185(2–4) 153-164.
- Villamor, G.B., Le, Q.B., Djanibekov, U., van Noordwijk, M., Vlek, P.L.G., 2014. Biodiversity in rubber agroforests, carbon emissions, and rural livelihoods: An agent-based model of land-use dynamics in lowland Sumatra. *Environmental Modelling & Software* 61 151-165.
- Voinov, A., Bousquet, F., 2010. Modelling with stakeholders. *Environmental Modelling & Software* 25(11) 1268-1281.
- Wainwright, J., 2008. Can modelling enable us to understand the rôle of humans in landscape evolution? *Geoforum* 39(2) 659-674.
- Wilensky, U., Rand, W., 2015. *An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo*. MIT Press.
- Zhang, H., Li, Y.-F., Tan, H.B.K., 2010. Measuring design complexity of semantic web ontologies. *Journal of Systems and Software* 83(5) 803-814.